Discrete Mathematics 67 (1987) 177–189
North-Holland

ALTERNATING WHITNEY SUMS AND MATCHINGS IN TREES, PART 1

Robert E. JAMISON*
Department of Mathematical Sciences, Clemson University, Clemson, SC 29634-1907, U.S.A.

Rececived 2 October 1984
Revised 10 June 1986

The number of k-node subtrees of a tree is its kth Whitney number. This paper investigates the behavior of certain alternating sums of these Whitney numbers and shows how they are related to the structure of maximum matchings in the tree. It is shown that the alternating sum of the Whitney numbers gives the maximum cardinality of an independent set of nodes. Moreover, a weighted alternating sum yields the number of nodes left uncovered by at least one maximum matching.

1. Introduction

Throughout this paper T will denote a tree (connected, acyclic, undirected graph) on $|T| = n$ nodes. By a *k-subtree S* of T we shall mean a set of k nodes of T which induce a connected subgraph of T. The collection of all subtrees of T forms a meet-distributive lattice [1, 2] in which the k-subtrees are precisely the elements of height (or rank) k. The number $A_k(T)$ of k-subtrees is thus the kth *Whitney number* (or rank number) [4] of the lattice of subtrees. The Whitney numbers for some selected trees (Fig. 1) are given in Table 1.

For a few values of k, the Whitney numbers are given by fairly simple, explicit formulae:

$$A_1(T) = n, \quad \text{the number of nodes of } T,$$

$$A_2(T) = n - 1, \quad \text{the number of edges of } T,$$

$$A_3(T) = \sum_{d=2} V_d(T)\binom{d}{2},$$

$$A_{n-2}(T) = \binom{V_1(T)}{2} + V_1^*(T),$$

$$A_{n-1}(T) = V_1(T), \quad \text{the number of endnodes of } T,$$

$$A_n(T) = 1.$$

Here $V_d(T)$ denotes the number of vertices of degree d in T and $V_1^*(T)$ denotes

* The author was supported for this research from NSF Grant MCS-80-02543 and NSF EPSCOR Grant ISP-80-11451.

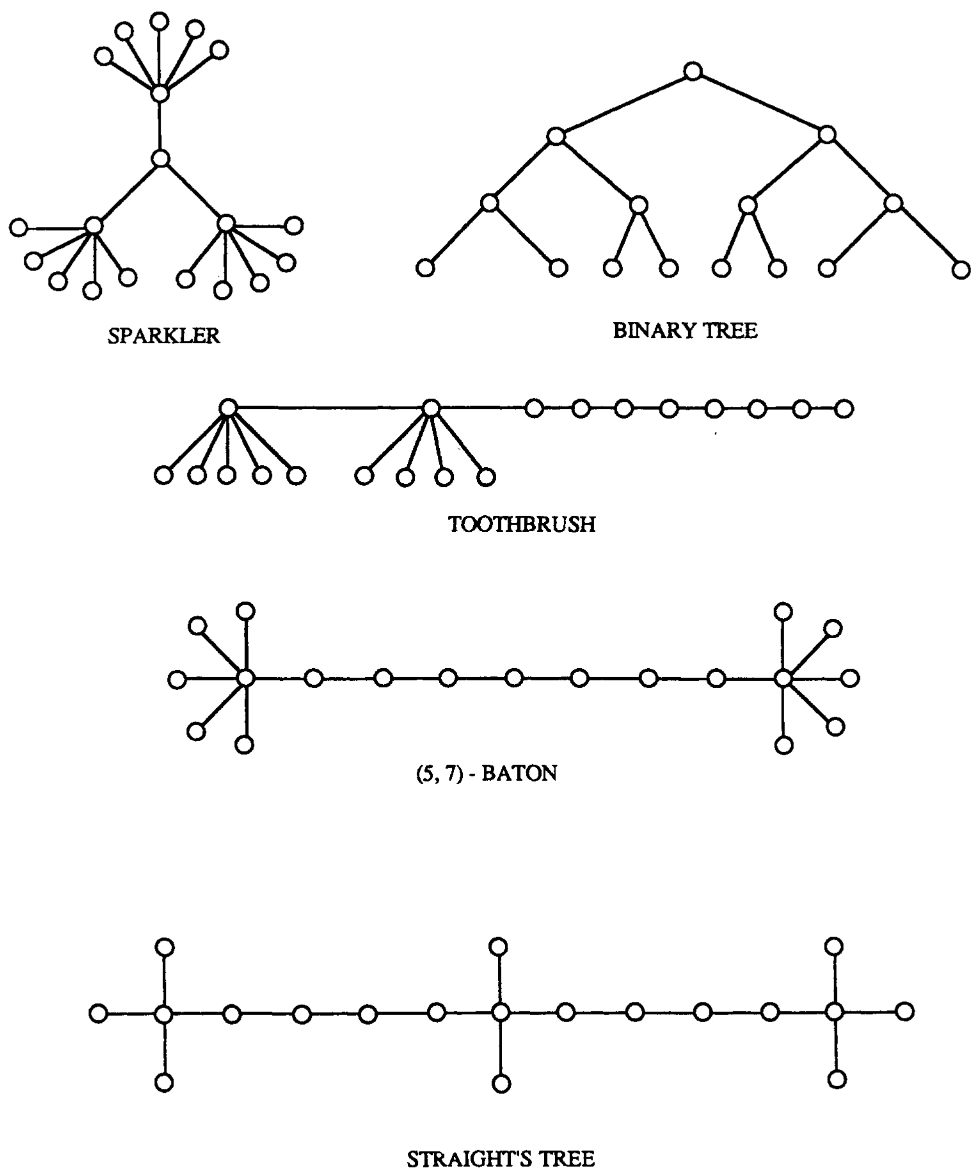

Fig. 1. Some selected trees.

the number of endnodes adjacent to nodes of degree 2. The verification of these formulae is straightforward and left to the reader, who should note that for other values of k, explicit formulae are increasingly complex and less informative. Fortunately the Whitney numbers of any given tree T can be calculated by a fairly simple recursive routine which will be developed in the next section (Theorem 2.1).

In [6] the extreme values of the Whitney numbers were shown to be attained by the path and the star and given by

$$n-k+1 \leqslant A_k(T) \leqslant \binom{n-1}{k-1}, \quad k>1.$$

Moreover, if for some k with $2<k<n$, the lower (upper) bound is achieved, then T must be a path (star). Beyond this there is little known about the behavior of the individual A_k.

Table 1. Whitney numbers of selected trees. (See Fig. 1.)

k	Sparkler	Toothbrush	(5, 7)-Baton	Straight's tree	Binary tree
1	19	19	19	19	15
2	18	18	18	18	14
3	48	37	37	26	19
4	91	76	56	30	26
5	195	155	65	32	38
6	483	276	66*	34	52
7	1088	400	65	36	71
8	2121	483	64	46	94
9	3633	518	63	60	114
10	5365	526*	62	66*	116*
11	6570*	522	77	64	94
12	6465	507	132	60	60
13	5008	467	212	60	28
14	3003	382	252*	68	8
15	1365	256	210	72*	1
16	455	130	120	56	
17	105	46	45	28	
18	15	10	10	8	
19	1	1	1	1	
M_T	11.16	10.28	11.44	10.47	8.51

* indicates a mode.

Regarding the behavior of the sequence $A_k(T)$ as a whole one might guess from small order examples (say, $n \leq 15$) that this sequence is unimodal. (We ignore here the initial dip from $A_1(T) = n$ to $A_2(T) = n - 1$ that occurs in all trees.) Table 1 reveals, however, two trees on 19 nodes that have two nontrivial modes (marked by *). David Whited [13] has even discovered trees with as many as 6 nontrivial modes. (The 6 mode example is a caterpillar with $n = 354$ nodes.) Although the pattern suggested by Whited's examples fails for 7 modes, it seems likely that there are trees with arbitrarily many modes. All known multi-modal trees have many vertices of degree 2, and it is tempting to conjecture that any homeomorphically irreducible tree (i.e., $V_2(T) = 0$) has unimodal Whitney numbers.

In spite of the considerable variation in the individual Whitney numbers, it turns out that certain alternating sums have a very regular behavior:

$$E_T = \sum_{k=1}^{n} (-1)^{k-1} A_k(T), \tag{1.1}$$

$$H_T = \sum_{k=1}^{n} (-1)^{k-1} k A_k(T), \tag{1.2}$$

$$I_T = \sum_{k=1}^{n} (-1)^{k-1} k^2 A_k(T), \tag{1.3}$$

Without the alternating signs, the first sum would count the subtrees of T and the second would be the sum of their cardinalities. Their ratio M_T, the average order of a subtree of T, has been studied in [6, 7 and 9].

With the alternating signs, the sum E_T is the excess in the number of odd order subtrees over the even order subtrees. Our goal here is to establish (at the end of Section 3) the following less obvious combinatorial interpretations of the sums E_T and H_T.

Theorem 1.4. *For any tree T, E_T is the maximum cardinality of an independent set of nodes of T, and H_T is the cardinality of the intersection of all maximum independent sets.*

This immediately implies that $0 \leqslant H_T \leqslant E_T \leqslant |T|$. We will also establish (Theorem 4.5) the lower bound $2E_T - |T| + 1 \leqslant H_T$ if $H_T > 1$, and show (Theorem 4.6) that all possibilities for E_T, H_T, and $|T|$ satisfying these bounds actually occur. The average value of E_T (as the maximum cardinality of an independent set of nodes) has been investigated by Meir and Moon [8].

The invariant I_T does not seem to have a simple combinatorial interpretation. However, in Part 2 the bounds below will be established:

$$-\tfrac{1}{4}(n^2 + 2n) < I_T < \tfrac{1}{4}(n+1)^2. \tag{1.5}$$

It will also be shown in Part 2 that the lower bound is achieved iff T is a path on an even number of nodes, and the upper bound is achieved iff T is an alternating tree in the sense of Edmonds [3].

2. Recursive calculation of the Whitney numbers and their sums

It is useful to consider the generating function (the "rank polynomial" of T)

$$\Phi_T(x) = \sum_{k=1}^{n} A_k(T)x^k,$$

for the Whitney numbers of a tree T. As in [6], it is also helpful to consider a local version as well. If p is any node of T, let

$$\varphi_T(p;x) = \sum_{k=1}^{n} \alpha_k(T;p)x^k,$$

where $\alpha_k(T;p)$ is the number of k-subtrees of T which contain p. The notation T_p will refer to the tree T rooted at p. Let

$$D(v; T_p) = \{w \in T\colon v \text{ lies on the path from } w \text{ to } p\},$$

denote the set of *descendents* of v in the rooted tree T_p. For convenience of notation, $\varphi_T(v \mid p;x)$ will denote $\varphi_D(v;x)$, where $D = D(v; T_p)$. The *children* of

v in T_p are the neighbors of v in $D(v; T_p)$. When it is clear from context, the subscript T may be dropped from φ_T.

Theorem 2.1. *For any node p in a tree T,*

(a) $\varphi(p; x) = x \prod (1 + \varphi(v \mid p; x))$, *where the product runs over all neighbors v of p in T, and*

(b) $\Phi_T(x) = \sum \varphi(v \mid p; x)$, *where the sum runs over all nodes v of T.*

Proof. (a) Let $v_1, \ldots, v_d$ be the neighbors of p, and let $B_i = D(v_i, T_p)$ denote the branch of T at p containing v_i. Any subtree through p is formed by joining together subtrees from the branches B_i. Since the component subtree in the ith branch either is empty or must contain v_i, the recursion follows.

(b) Rooting T at p will partially order the nodes of T by the relation $w \leqslant v$ iff v lies on the path from w to p—i.e., iff w is a descendent of v. In this order, T_p is a join-semilattice—that is, every set of nodes has an supremum.

Now suppose S is a subtree of T and v is a node of T. If $v > \sup S$, then $v \notin S$. If $v < \sup S$ or if v and $\sup S$ are not comparable, then S is not contained in $D(v; T_p)$. Hence in each of these cases, S is not counted by $\varphi(v \mid p; x)$. But if $v = \sup S$, then $v \in S$ and $S \subseteq D(v; T_p)$. Whence S is counted by $\varphi(v \mid p; x)$ precisely for $v = \sup S$. Thus each subtree of T is counted exactly once by the sum in Theorem 2.1(b). $\square$

The case $x = 1$ of Theorem 2.1(a) was used by Ruskey [10] in the study of the average number of rooted subtrees in a random planar tree. The general case of Theorem 2.1(a) was exploited in [6] in the study of the average order of subtrees of a tree. Together (a) and (b) of Theorem 2.1 yield an easily implemented algorithm for finding the Whitney numbers of a tree T. Namely, root T at any node p, and recursively attach polynomial labels $\varphi(v \mid p; x)$ to the nodes of T as follows: First label all leaves (endnodes except p) with x. Then when all the children of a node v are labelled, the label $\varphi(v \mid p; x)$ on v may be calculated by applying Theorem 2.1(a) to v in $D(v; T_p)$. Finally by Theorem 2.1(b), $\Phi_T(x)$ is the sum of the resulting labels. This procedure is illustrated in Fig. 2. An implementation of this algorithm by David Whited was used to obtain the entries in Table 1.

Note that the invariants (1.1) and (1.2) may be obtained from the rank polynomial Φ_T and its derivative Φ_T' by taking $x = -1$:

$$E_T = -\Phi_T(-1) \quad \text{and} \quad H_T = \Phi_T'(-1).$$

Their local analogues, which will play an important role, are given by

$$\varepsilon_T(p) = -\varphi_T(p; -1) \quad \text{and} \quad \eta_T(p) = \varphi_T'(p; -1).$$

The relative values $\varepsilon_T(v \mid p)$ and $\eta_T(v \mid p)$ are defined similarly from $\varphi_T(v \mid p; x)$. Again the subscript T may be dropped when it is clear from context.

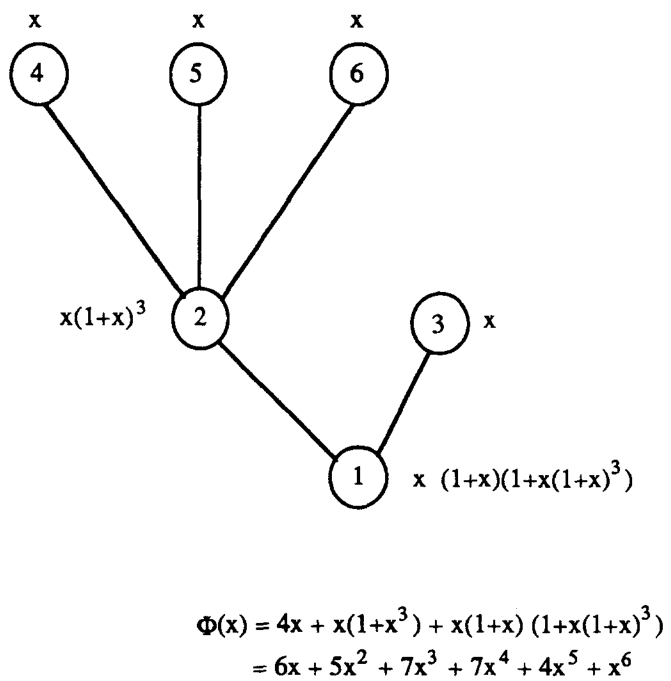

Fig. 2. Recursive determination of the Whitney numbers.

Setting $x = -1$ in Theorem 2.1(a) yields

$$\varepsilon_T(p) = \prod (1 - \varepsilon(v \mid p)), \tag{2.2}$$

where the product runs over all neighbors v of p. Since $\varepsilon_T(p) = 1$ if $|T| = 1$, the following may be obtained by induction on $|T|$.

Theorem 2.3. *For any node p in any tree T, $\varepsilon_T(p)$ is either* 0 *or* 1.

Setting $x = -1$ into Theorem 2.1(b) and its derivative yields

$$E_T = \sum_{v \in T} \varepsilon(v \mid p) \quad \text{and} \quad H_T = \sum_{v \in T} \eta(v \mid p). \tag{2.4}$$

An expression for H_T may also be obtained in terms of $\varepsilon_T(p)$. Namely, note that

$$x\Phi'_T(x) = \sum_{k=1}^{n} kA_k(T)x^k = \sum_{p \in T} \varphi_T(p; x), \tag{2.5}$$

since in the last sum each k-subtree contributes a term x^k for each of its points. Setting $x = -1$, gives

$$H_T = \sum_{p \in T} \varepsilon_T(p). \tag{2.6}$$

Differentiating (2.5) and setting $x = -1$ yields

$$I_T = \sum_{p \in T} \eta_T(p). \tag{2.7}$$

We now derive a direct recursion for ε and η which will allow us to conclude that ε is, in fact, determined by η. Differentiating the equation of Theorem 2.1(a) and setting $x = -1$ yields

$$\eta(p) = \prod_{i=1}^{d} (1 - \varepsilon(v_i \mid p)) - \sum_{i=1}^{d} \left[\eta(v_i \mid p) \prod_{j \neq i} (1 - \varepsilon(v_j \mid p)) \right], \tag{2.8}$$

where $v_1, \ldots, v_d$ are the neighbors of p. By Theorem 2.3 each term in the above products is either 0 or 1. This leads, along with (2.2), to simple recursive rules for the determination of ε and η:

If $\varepsilon(v \mid p) = 1$ for two or more neighbors v of p, then $\varepsilon(p) = 0$ and $\eta(p) = 0$, (2.9*a*)

If $\varepsilon(v \mid p) = 1$ for exactly one neighbor v of p, then $\varepsilon(p) = 0$ and $\eta(p) = -\eta(v \mid p)$, (2.9b)

If $\varepsilon(v \mid p) = 0$ for all neighbors v of p, then $\varepsilon(p) = 1$ and $\eta(p) = 1 - \sum \eta(v \mid p)$, (2.9c)

where the sum ranges over all neighbors v of p.

For any node p in T, the values $\varepsilon_T(p)$ and $\eta_T(p)$ may now be easily computed recursively using the rules (2.9). Take two copies of the tree T rooted at p. One will be labelled with the values $\varepsilon_T(v \mid p)$, the other with $\eta_T(v \mid p)$, by starting at the leaves and working toward the root. If $v \neq p$ is a leaf, then $D(v; T_p)$ consists of v alone, so $\varepsilon(v \mid p) = \eta(v \mid p) = 1$. When all children of a node q are labelled, then q may also be labelled applying the rules (2.9) in the subtree $D(q; T_p)$. Finally the values $\varepsilon_T(p) = \varepsilon(p \mid p)$ and $\eta_T(p) = \eta(p \mid p)$ will be obtained. This process is illustrated in Fig. 3.

Starting with the case $\varepsilon_T(p) = \eta_T(p) = 1$ if T consists of a single node p, one

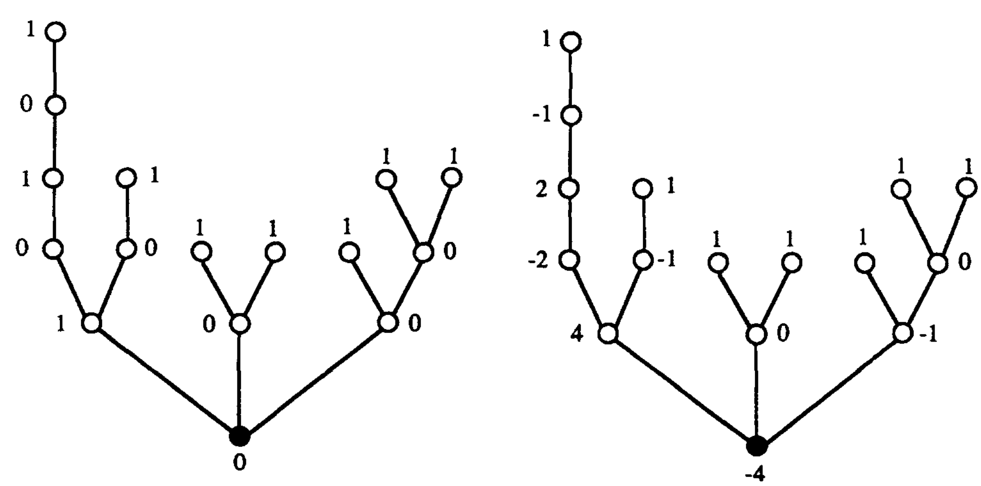

Fig. 3. Recursive calculation of ε and η in a rooted tree.

can readily establish the following connection by induction on $|T| = n$ using (2.9):

Theorem 2.10. *For any tree T and node p in T, $\varepsilon_T(p) = 0$ iff $\eta_T(p) \leqslant 0$ and $\varepsilon_T(p) = 1$ iff $\eta_T(p) \geqslant 1$.*

It follows that in the above procedure it is possible, and in fact sufficient, to work with the η-labelled tree alone.

3. p-Matchings

Let G be an arbitrary graph. A *matching* (cf. [12]) is a set $\mathcal{M}$ of vertex-disjoint edges in G. A node is *covered* by $\mathcal{M}$ if it is an endnode of an edge in $\mathcal{M}$. The *length* of $\mathcal{M}$ is the number of edges in $\mathcal{M}$, and the *defect* of $\mathcal{M}$ is the number of nodes of G not covered by $\mathcal{M}$. The maximum length [minimum defect] over all matchings is the *length* $\lambda(G)$ (*defect* $\delta(G)$) *of* G. A matching of maximum length (and hence minimum defect) is a *maximum matching*. (*Warning*: A matching may be maxim*al* in inclusion but not maxim*um* in length—e.g., the middle edge of a path on 4 nodes.) A *perfect matching* (1-factor) covers all nodes of G (i.e., has defect 0).

A set I nodes of G is *independent* if there are no edges between nodes in I. Let $\mu(G)$ denote the maximum cardinality of an independent set in G. An independent set I of cardinality $|I| = \mu(G)$ is a *maximum independent set.* Again, an independent set may be maxim*al* in inclusion but not maxim*um* in cardinality—e.g., the second and fourth nodes in a path on 5 nodes.

If $\mathcal{M}$ is a matching and I an independent set, then each edge of $\mathcal{M}$ must contain at least one node in the complement $G \backslash I$ of I. Thus the length $|\mathcal{M}|$ is at most $|G \backslash I|$. Hence for any graph G on n nodes:

$$\lambda(G) + \mu(G) \leqslant n. \tag{3.1}$$

Now let T be a tree rooted at a node p. Suppose v is a node of T with $\varepsilon(v \mid p) = 0$. Then by (2.9) applied to the branch $B = D(v; T_p)$, there is a child v' of v with $\varepsilon_B(v' \mid v) = 1$. Since the branches $D(v'; B_v)$ and $D(v'; T_p)$ obviously coincide, $\varepsilon_T(v' \mid p) = \varepsilon_B(v' \mid v) = 1$. Thus for each v with $\varepsilon_T(v \mid p) = 0$, we may select a child v' of v with $\varepsilon_T(v' \mid p) = 1$. It is possible that this selection can be done in several ways. However, since v is the unique parent of v' in the rooted tree T, we never can have $w' = v'$ for any $w \neq v$. Hence the edges vv' are vertex-disjoint and thus form a matching. Any matching obtained in this way will be called a *p-matching* of T. Matchings of this type have been used by Carla Savage [11] to give a linear time algorithm for finding a maximum matching in a tree.

Theorem 3.2. *Let p be any node in a tree T,*

(a) *Every p-matching is maximum,*

(b) *The set $I(T_p) = \{v \in T : \varepsilon_T(v \mid p) = 1\}$ is a maximum independent set in T.*

Proof. Let us first note that $I(T_p)$ is independent. Indeed, suppose vw is an edge of T. Then regarded from the root p, one of v and w is the parent and the other the child. It follows from the rules (2.9) that if the child belongs to $I(T_p)$, then the parent cannot. Hence v and w cannot both be in $I(T_p)$.

Now let $\mathcal{M}$ be any p-matching. By definition, there is an edge of $\mathcal{M}$ for each node v of T with $\varepsilon_T(v \mid p) = 0$. From this and (3.1) we get

$$n = |\mathcal{M}| + |I(T_p)| \leqslant \lambda(G) + \mu(G) \leqslant n, \tag{3.3}$$

so equality holds throughout. This forces $|\mathcal{M}| = \lambda(G)$ and $|I(T_p)| = \mu(G)$, whence (a) and (b) follow. □

It is worth noting that the "converse" of Theorem 3.2 does not hold in general. Indeed, suppose w is an endnode of T adjacent to a node v of degree 2. If $p \notin \{v, w\}$, then w is the unique child of v in T_p. Now $\varepsilon(w \mid p) = 1$ since w is a leaf, so $\varepsilon(v \mid p) = 0$ by (2.9). It follows that any p-matching must match v with w. Using this, one can see that the maximum matching in Fig. 4 is not a p-matching for any p. Likewise, $I(T_p)$ must contain all endnodes unless p is an endnode, in which case $I(T_p)$ must contain all endnodes except possibly p. Thus the maximum independent set in Fig. 5 is not of the form $I(T_p)$ for any p.

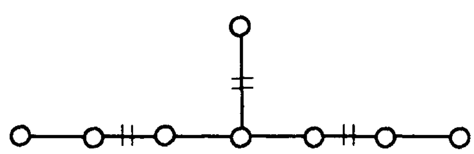

Fig. 4. Maximum matching that is not a p-matching for any p.

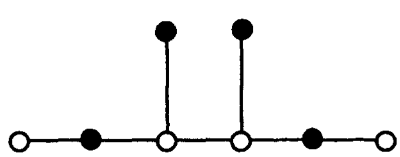

Fig. 5. Maximum independent set that is not of the form $I(T_p)$ for any p.

For any arbitrary graph G, let G^+ denote the set of all nodes of G which are left uncovered by at least one maximum matching in G.

It follows from (3.3) that equality holds in (3.1) for any tree. As a consequence of König's celebrated theorem on matchings, this is true for arbitrary bipartite graphs. This implies a characterization of G^+ which will be useful below.

Lemma 3.4. *Let G be a bipartite graph with n nodes,*

(a) $\lambda(G) + \mu(G) = n$,

(b) *G^+ is the intersection of all maximum independent sets in G.*

Proof. (a) is a well-known consequence of the theorems of Gallai and König [5].

(b) Let $\mathcal{M}$ be any maximum matching of G and I any maximum independent set. By (a),

$$|G\setminus I| = n - \mu(G) = \lambda(G) = |\mathcal{M}|,$$

so there are the same number of nodes in the complement $G\setminus I$ of I as there are edges in $\mathcal{M}$. Since no edge of $\mathcal{M}$ can have both endnodes in I, it follows that $\mathcal{M}$ matches $G\setminus I$ into I.

Thus if p is left uncovered by $\mathcal{M}$, then p must belong to I. Hence G^+ is contained in the intersection of all maximum independent sets.

Conversely, suppose $p \notin G^+$. Then p is covered by every maximum matching. Say pq is the edge of $\mathcal{M}$ covering p. Let $\mathcal{M}'$ be the matching of $G\setminus p$ obtained by deleting pq from $\mathcal{M}$. If $\mathcal{M}'$ is not maximum in $G\setminus p$, there would be a matching of length $|\mathcal{M}'| + 1 = \lambda(G)$ in $G\setminus p$. But this would be a maximum matching of G missing p, contradicting $p \notin G^+$. Thus $\mathcal{M}'$ is maximum in $G\setminus p$, so $\lambda(G\setminus p) = \lambda(G) - 1$. By (a), $\mu(G\setminus p) = (n-1) - \lambda(G\setminus p) = n - \lambda(G) = \mu(G)$. Thus if J is any maximum independent set of $G\setminus p$, then J is also a maximum independent set of G. Now q is uncovered by $\mathcal{M}'$, so by the previous paragraph applied to $G\setminus p$, q belongs to J. But then J is a maximum independent set of G missing p, whence b) is established. □

Lemma 3.5. *For any tree T, $T^+ = \{p \in T\colon \varepsilon_T(p) = 1\}$.*

Proof. Suppose $\varepsilon_T(p) = 1$. Then $\varepsilon_T(p \mid p) = \varepsilon_T(p) = 1$, so by construction no p-matching will cover p. Since p-matchings are maximum by Theorem 3.2, it follows that $p \in T^+$. Conversely, if $p \in T^+$, then since $I(T_p)$ is maximum by Theorem 3.2, it follows from Lemma 3.4 that $p \in I(T_p)$. Thus $\varepsilon_T(p) = \varepsilon_T(p \mid p) = 1$. □

It is now easy to establish the interpretations of E_T and H_T stated in the Introduction.

Proof of Theorem 1.4. By Theorem 2.3 and (2.4), E_T is just the cardinality of $I(T_p)$, which by (3.2) is a maximum independent set. By (2.6), Lemma 3.5 and Lemma 3.4, H_T is the number of nodes which lie in all maximum independent sets. □

4. The range of values for E_T and H_T

To conclude, we determine the range of possible values for the invariants E_T and H_T. Sharp bounds on E_T and H_T may be obtained from quite general considerations via the interpretations in Theorem 1.4.

Lemma 4.1. *Suppose G is a connected graph with more than one node. If G does not have a perfect matching, then $|G^+| \geqslant \delta(G)+1$.*

Proof. Let $\mathcal{M}$ be a maximum matching. The $\delta(G)>0$ nodes left uncovered by $\mathcal{M}$ are all in G^+. Let p be such a node. Since G is connected and p is not its only node, p has a neighbor v. Clearly $\mathcal{M}$ must cover v or the edge vp could be added to $\mathcal{M}$. Say the edge vw is in $\mathcal{M}$. Then replacing this edge by vp results in a maximum matching missing w. Thus w is also in G^+, so $|G^+| \geqslant \delta(G)+1$. □

Lemma 4.2. *If G is a connected graph and $|G^+|=1$, then G consists of a single node.*

Proof. The fact that $G^+ \neq \emptyset$ means G does not have a perfect matching. If G had more than one point, then by Lemma 4.1 we would have $|G^+| \geqslant \delta(G)+1 \geqslant 2$, a contradiction. □

Lemma 4.3. *If G is a connected bipartite graph on n nodes with $|G^+|>1$, then*

$$\tfrac{1}{2}n < \mu(G) \leqslant n-1 \quad \textit{and} \quad 2\mu(G)-n+1 \leqslant |G^+| \leqslant \mu(G).$$

Proof. Clearly $2\lambda(G)+\delta(G)=n$. By König's Theorem (Lemma 3.4(a)), $\lambda(G)+\mu(G)=n$. Thus

$$\mu(G)=n-\lambda(G)=\lambda(G)+\delta(G)=\tfrac{1}{2}(n+\delta(G)). \tag{4.4}$$

Since $G^+ \neq \emptyset$, G does not have a perfect matching, so $\delta(G)>0$. Thus $\mu(G)>\frac{1}{2}n$ follows. The bound $\mu(G) \leqslant n-1$ is obvious since G has more than one point.

Now by (4.4), $\delta(G)=2\mu(G)-n$. Hence $|G^+| \geqslant 2\mu(G)-n+1$ follows from Lemma 4.1. The inequality $|G^+| \leqslant \mu(G)$ follows from Lemma 3.4(b). □

Theorem 4.5. *Let T be any tree on n nodes,*

(a) *$H_T=0$ iff T has a perfect matching,*
(b) *$H_T=1$ iff T has just one node,*
(c) *If $H_T>1$, then $\frac{1}{2}n<E_T \leqslant n-1$ and $2E_T-n+1 \leqslant H_T \leqslant E_T$.*

Proof. By Theorem 1.4, $E_T=\mu(G)$ and $H_T=|T^+|$. Thus (a) follows since $T^+=\emptyset$ means that no nodes are left uncovered by any maximum matching. Also (b) follows from Lemma 4.2 and (c) follows from Lemma 4.3. □

Theorem 4.6. *Suppose h, e, and n are positive integers such that*

$$(1)\ \tfrac{1}{2}n<e \leqslant n-1, \quad \textit{and} \quad (2)\ 2e-n+1 \leqslant h \leqslant e.$$

Then there is a tree T with n nodes such that $E_T=e$ and $H_T=h$.

Proof. We proceed by induction on n. The smallest case is $n = 3$ with $e = h = 2$, which corresponds to the path on 3 nodes. Now in general, if $e = n - 1$, then also $h = n - 1$ by (2), and the star $T = K_{1,n-1}$ is the desired tree. Thus suppose hereafter that $e \leqslant n - 2$. Set $e' = e - 1$ and $n' = n - 2$. The relation $\frac{1}{2}n' < e' \leqslant n' - 1$ then holds.

Case 1. $h = 2e - n + 1$

We then have $2e' - n' + 1 = 2e - n + 1 = h \leqslant e + (n-2) - n + 1 = e'$, so h, e', and n' satisfy (1) and (2). Thus there is a tree S with $|S| = n'$ and $E_S = e'$ and $H_S = h$. Now $H_S \leqslant e' \leqslant n' - 1$, so there is at least one node p in S not in S^+. Attach two new nodes v and w to S to form a tree T by making v adjacent to p and w adjacent to v. Then $|T| = n$. If I is a maximum independent set in S, then $I \cup w$ is a maximum independent set in T. Thus $E_T = e$. Now any maximum matching $\mathcal{M}$ of S covers p since $p \neq S^+$. Adding the edge vw to $\mathcal{M}$ results in a maximum matching in T, and all such arise this way. It follows that $T^+ = S^+$, so $H_T = H_S = h$. Thus T is the required tree.

Case 2. $h > 2e - n + 1$

Letting $h' = h - 1$, we have $2e' - n' + 1 = 2e - n + 1 \leqslant h - 1 = h' \leqslant e - 1 = e'$, so h', e' and n' satisfy (1) and (2). Thus there is a tree S with $|S| = n'$ and $E_S = e'$ and $H_S = h'$. Now $H_S \geqslant 2e - n + 1 > 1$, by (1), so S^+ is nonempty. Let p be a node in S^+ and attach new nodes v and w as above to obtain a new tree T with $|T| = n$ and $E_T = e$. Now since $p \in S^+$, there is a maximum matching $\mathcal{M}$ of S which leaves p uncovered. Adding the edge pv to $\mathcal{M}$ results in a maximum matching of T with w uncovered. Thus $w \in T^+$. Since T^+ is independent by (3.4), $v \neq T^+$. If follows that $T^+ = S^+ \cup w$. Thus $H_T = H_S + 1 = h$. Whence T is the desired tree. □

Acknowledgment

The author is deeply grateful to J.W. Moon whose insightful comments on a preliminary version of this paper resulted in many improvements.

References

[1] P.H. Edelman, Meet-distributive lattices and the anti-exchange closure, Algebra Universalis 10 (1980) 290–299.

[2] P.H. Edelman and R.E. Jamison, The theory of convex geometrics, Geom. Dedicata 19 (1985) 247–270.

[3] J. Edmonds, Paths, trees, and flowers, Canad. J. Math. 17 (1965) 449–467.

[4] C. Greene and D.J. Kleitman, Proof techniques in the theory of finite sets, in: G.C. Rota, ed., Studies in Combinatorics, MAA Studies 17 (1978) 22–79.

[5] F. Harary, Graph Theory (Addison-Wesley, Reading, MA, 1972) 95–96.
[6] R.E. Jamison, On the average number of nodes in a subtree of a tree, J. Combin. Theory Ser. B 35 (1983) 207–223.
[7] R.E. Jamison, Monotonicity of the mean order of subtrees, J. Combin. Theory Ser. B 37 (1984) 70–78.
[8] A. Meir and J.W. Moon, The expected node-independence number of random trees, Indag. Math. 35 (1974) 335–341.
[9] A. Meir and J.W. Moon, On subtrees of certain families of rooted trees, Ars Combin. 16-B (1983) 305–318.
[10] F. Ruskey, Listing and counting subtrees of a tree, SIAM J. Comput. 10 (1981) 141–150.
[11] C. Savage, Maximum matchings and trees, Inform. Process. Letters 10 (1980) 202–205.
[12] D.J.A. Welsh, Matroid Theory (Academic Press, London, 1976) 97–115.
[13] D. Whited, private communication (1981).